DOCUMENT RESUME

ED 392 806 TM 024 395

AUTHOR	Gilford, Dorothy M., Ed.
TITLE	A Collaborative Agenda for Improving International Comparative Studies in Education.
INSTITUTION	National Academy of Sciences - National Research Council, Washington, DC. Board on International Comparative Studies in Education.
SPONS AGENCY	National Center for Education Statistics (ED), Washington, DC.; National Science Foundation, Washington, D.C.
PUB DATE	93
NOTE	51p.
AVAILABLE FROM	Board on International Comparative Studies in Education, National Research Council, 2101 Constitution Avenue, N.W., HA 178, Washington, DC 20418.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Agenda Setting; Comparative Education; *Cooperation; Cross Cultural Studies; *Data Collection; Educational Planning; *Educational Research; Elementary Secondary Education; Information Dissemination; International Education; *International Studies; Policy Formation; Systems Development
IDENTIFIERS	*Board on International Comparative Studies Educ

ABSTRACT

The Board on International Comparative Studies in Education monitors U.S. participation in cross-national comparative studies in education and provides advice to sponsoring agencies. This document results from the Board's consideration of studies that would enrich plans for international studies. What the Board believes to be the value of international comparative studies and how to improve such studies are discussed. Establishing a more coherent and effective worldwide system for research and data collection in education is a priority that will require common indicators for describing educational accomplishments, research and syntheses of research, documentation of practices, dissemination of information, and archives and research reports. Establishing a data collection, reporting, and dissemination system must take into account the views of interested parties, the technical and substantive state of the art, and the promise of cross-national studies for responding to policy needs and advancing knowledge. (Contains 18 references.) (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 392 806

A Collaborative Agenda for Improving International Comparative Studies in Education

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

JANET S. HANSEN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

BEST COPY AVAILABLE

11902345

A Collaborative Agenda for Improving International Comparative Studies in Education

Dorothy M. Gilford, Editor

Board on International Comparative Studies in Education

Michael W. Kirst, Chair

Commission on Behavioral and Social Sciences and Education

National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C. 1993

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

This project was supported with funds from the National Science Foundation and the National Center for Education Statistics, U.S. Department of Education.

Additional copies of this report are available from:

Board on International Comparative Studies in Education
National Research Council
2101 Constitution Avenue, N.W. HA 178
Washington, D.C. 20418

Printed in the United States of America

Copyright 1993 by the National Academy of Sciences. All rights reserved.

BOARD ON INTERNATIONAL COMPARATIVE STUDIES IN EDUCATION

MICHAEL W. KIRST (*Chair*), School of Education, Stanford University
JOSEPH F. ALIBRANDI,* Whittaker Corporation, Los Angeles, California
GORDON M. AMBACH (*ex-officio*), Council of Chief State School Officers, Washington, D.C.
LLOYD BOND,** Center for Educational Research and Evaluation, University of North Carolina
HERMAN CHERNOFF,* Department of Statistics, Harvard University
TIMOTHY H. GOLDSMITH, Department of Biology, Yale University
JAMES W. GUTHRIE, School of Education, University of California
EDWARD H. HAERTEL, School of Education, Stanford University
STEPHEN P. HEYNEMAN, Human Resources Technical Department, Europe and Central Asia and Mid East and N. Africa Regions, The World Bank, Washington, D.C.
DANIEL G. HORVITZ, Research Triangle Institute, Research Triangle Park, North Carolina
GAEA LEINHARDT,* Department of Education, University of Pittsburgh
PAUL G. LEMAHIEU,** University of Delaware and Delaware Department of Public Instruction
MARY M. LINDQUIST, School of Education, Columbus College
HAROLD NISSELSON,** independent statistical consultant, Falls Church, Virginia
ANDREW C. PORTER,** Wisconsin Center for Education Research, University of Wisconsin
FRANCISCO O. RAMIREZ, School of Education, Stanford University

JOHN R. SCHWILLE, College of Education, Michigan State University
MARSHALL S. SMITH (*former Chair*),* School of Education, Stanford University
FLORALINE I. STEVENS,* Los Angeles Unified School District, Los Angeles, California
JUDITH V. TORNEY-PURTA, Department of Human Development, University of Maryland

DOROTHY M. GILFORD, Staff Director
M. JANE PHILLIPS, Administrative Assistant

*Former member who contributed to this publication.
**As a new member, has not participated in the writing or review of this report.

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce Alberts is president of the National Academy of Sciences.

The National Academy of Engineering was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Robert M. White is president of the National Academy of Engineering.

The Institute of Medicine was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Kenneth I. Shine is president of the Institute of Medicine.

The National Research Council was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce Alberts and Dr. Robert M. White are chairman and vice chairman, respectively, of the National Research Council.

PREFACE

The Board on International Comparative Studies in Education was established in 1988 at the request of the U.S. National Center for Education Statistics and the U.S. National Science Foundation. The board operates under the aegis of the Commission on Behavioral and Social Sciences and Education of the National Research Council. The board monitors U.S. participation in cross-national comparative studies in education that are funded by its sponsors and provides advice to the sponsoring agencies. In carrying out its responsibilities, the board provides a forum for information and discussion; assists in planning the conduct and funding of international studies; establishes principles regarding the quality of study design, data collection and analysis procedures, and report preparation; encourages dissemination of study findings; promotes the use of assessment findings to improve U.S. education; and periodically suggests studies that would be useful additions to agencies' plans for U.S. participation in international studies.

This document is the result of activities of the board in considering studies that would enrich the agencies' plans. In assessing the feasibility of a new international study, it is essential to consider the interest and willingness of other countries to participate in the study. At present, worldwide arrangements for planning and conducting cross-national research and data collection in education are inadequate. Responsibility is fragmented among many organizations in many countries and, consequently, there is no easy way to assess the feasibility of a study. Since reliable information based on sound cross-national research and data collection can be useful in improving educational practices and outcomes in the United States as well as in other nations, members of the board believe it is important to establish a coherent and effective worldwide system to develop such information.

This document includes discussion of what the board perceives to be the value of international comparative studies in education and its views on how these studies could be improved. It also includes a challenge to governments and lead-

ers in education to form alliances to develop the information needed to improve educational outcomes and practices; it also suggests a possible approach to meet the challenge. The board hopes that this document will encourage interest and a willingness to engage in this work by suggesting a vision of what international comparative research and data collection might produce in the next 20 years.

The board began with several discussions of this vision and of current international data collections. To serve as a further catalyst in encouraging the development of a worldwide system for planning and conducting cross-national research and data collection in education, the board convened an International Conference on Long-Range Planning for Large-Scale Collections of International Education Statistics in Paris in June 1992. Participants included representatives of intergovernmental, governmental agencies, and nongovernmental organizations. An earlier version of this document was prepared as a basis for discussions at the conference. Key topics for discussion included agenda items for cross-national research for the next 10-20 years, improvement of infrastructures for cross-national research, and cooperative efforts that might be taken to meet the data requirements for the next decade. The conferees agreed that it would be useful to meet again to discuss cooperation between organizations and strategies for the development of a consortium or a new international network for long-range planning.

Therefore, the board, in conjunction with the United Nations Educational, Scientific, and Cultural Organization (UNESCO) and the International Academy of Education (IAE), convened a second international conference in Washington in January 1993. At the second conference, key discussion topics included the status and problems of international education statistics as well as challenges and prospects for international cooperation in education statistics and research. Participants agreed to continue and further the cooperation that has begun with these two conferences. Colin Power, UNESCO assistant director general for education and UNESCO representative to the conferences, agreed that UNESCO would serve as the central organization for collecting information relating to these activities. Jacques Hallak, director, International Institute for Educational Plan-

ning, offered to host a third international conference in Paris in summer 1993. Several participants agreed to carry out various activities to prepare materials for the third meeting. With this second conference, the board concludes its role as a catalyst in stimulating international coordination for research and the collection of data, although it retains substantial and continuing hopes for the success of these enterprises.

This document reflects the collective thinking of the board. The first draft was the result of a hard day's work by three working groups comprising the entire board. The draft was then reworked by board members Judith Torney-Purta and John Schwille and staff director Dorothy Gilford. These people deserve a special note of thanks for turning several disparate pieces into a coherent draft. This draft was then discussed and further revised by the board and reviewed by over a dozen members of the education research community. I wish to thank all the board members who contributed to the stimulating discussions that led to the document as well as all the reviewers who took the time to send their reflective and cogent comments.

I am also grateful to Laura Lathrop, research assistant, and Jane Phillips, administrative assistant, who assisted Dorothy Gilford throughout the preparation of consecutive drafts of the document and to Christine McShane, editor of the Commission on Behavioral and Social Sciences and Education, whose highly professional editing contributed greatly to its readability.

Michael W. Kirst, *Chair*
Board on International Comparative
Studies in Education

CONTENTS

WHAT IS THE VALUE OF INTERNATIONAL COMPARATIVE STUDIES? 2

Responding to Particular Policy Needs, 4
Long-Term Development of Knowledge, 14

HOW CAN INTERNATIONAL COMPARATIVE STUDIES BE IMPROVED? 16

Research Design, Analysis, and Dissemination, 17
Methods for Assessing Current and Desired Outcomes of Education, 19
Comparability Across Nations, 20
Qualitative Studies and Large-Scale Surveys, 22
Improving Quality Control and Monitoring, 23
Fostering Accumulation, Synthesis, and Dissemination, 24

A CHALLENGE TO REALIZE THE VISION AND A POSSIBLE APPROACH 25

Improving the Infrastructure for Cross-National Research, 25
A Possible Approach, 26

REFERENCES 29

BIOGRAPHICAL SKETCHES 31

A COLLABORATIVE AGENDA FOR IMPROVING INTERNATIONAL COMPARATIVE STUDIES IN EDUCATION

Interest in international comparisons in education has increased rapidly over the past two decades, perhaps now exceeding that of the earlier era of public discussion in the late 1950s that was stimulated by the successful launching of Sputnik. This renewed interest offers both potential benefits and considerable risks. Earlier research has been cited in the press and in reports prepared for policy makers, but much of this public discussion has been selective, overly simplified, and a poor guide to public policy making. The increasing scrutiny of earlier studies has revealed their limitations and the consequent need for improvement in the planning, execution, and dissemination of international comparative research.

If the situation is to be improved, the tasks of establishing a more coherent and effective worldwide system for cross-national research and data collection in education must be shared among many organizations in many countries. This document does not presume to set the agenda for international research, or even to indicate the priorities for such research—activities that would be appropriately carried out only within the framework of a worldwide system for cross-national education studies. This document was prepared to stimulate interest and a willingness to engage in establishing such a system by suggesting a vision of what international comparative research and data collection might produce in the next 20 years.

By the year 2010, this vision calls for those with an interest in learning about education in other countries to have access to the following sources of data, information, and analysis:

- A set of widely accepted and widely understood indicators of the status and conditions of education in various countries;[1]

[1]Some members of the comparative education community question the usefulness of indicators. They question whether the benefits such indicators produce are worth the

- Syntheses of empirical research throughout the world, bringing such research to bear on broad comparative questions of wide interest;
- Specialized comparative monographs on questions of importance and interest to all educators;
- Documentation of teaching practices and school organization in a form that will be stimulating and useful for discussion and study by specialists and nonspecialists, including, for example, videotape archives of a wide range of classroom practice;
- Widely distributed publications of comparative information on issues of public policy, written in nontechnical language and responsive to the interests of journalists and the public;
- Routine translation, distribution, and use of primary documents, such as textbooks and syllabi, when these documents are sufficiently distinctive and important to be of interest to educators outside the country of origin; and
- Archives necessary to strengthen the methods used in cross-national studies, including both technical reports on what works methodologically and banks of instruments, items, and other applicable research elements and models for a variety of cross-national studies.

This document is primarily intended to focus attention on the nature of collaborative efforts to attain the vision. It includes discussion of the value of international comparative studies in education; how they could be improved; and a suggestion for consultation and forming alliances among potential collaborators.

WHAT IS THE VALUE OF INTERNATIONAL COMPARATIVE STUDIES?

Why should the United States or other countries be interested in supporting or participating in international education

cost, and they are concerned that the financial and political cost of producing truly comparable indicators exceeds the political willingness of many countries to subject their educational system to the potentially harsh comparison that would ensue.

studies? From the Board on International Comparative Studies in Education's viewpoint, the most important reason for U.S. participation in international studies of education is to improve understanding of our own education system (Bradburn and Gilford, 1990). Since the philosophy, organization, and practice of education varies across cultures and societies, a comparative perspective will increase our understanding of its many forms. Every aspect of education might benefit from comparison, from issues of teaching and learning, to matters of curriculum and organization, to broader issues of education policy. Since there are no absolute standards in education, comparison informs the standard we set for our children. Comparative studies also help policy makers monitor the success of education systems. And given that many people are reluctant to conduct controlled experiments with our children's education, comparison of natural variation is usually the most feasible way to study the effects of differing policies and practices.[2]

Nevertheless, much of education research is parochial, uninformed by the comparative knowledge that cross-national education research can provide. Researchers should be encouraged to develop broad and in-depth knowledge about cross-cultural similarities and differences and similarities in the way people learn. They should also undertake studies that examine the context of and participation in formal and informal education at all levels. Research is also needed to generate more narrowly focused knowledge that speaks to pressing problems of practice, such as identifying pedagogies and schools that best help children at risk to avoid becoming dropouts or knowing the role of algebra in curricula designed for today's students and how they learn it best.

It would be important, therefore, to seek ways to engage education researchers, practitioners, and consumers in the search

[2]Some critics have noted, however, that controlled experiments have the potential to improve the methodological sensitivity of education studies. Recommendations for greater use of experimental methods are discussed in a recent National Research Council report on bilingual education studies (Meyer and Fienberg, 1992). An exclusive reliance on correlational research designs and data analysis techniques is a methodologically limited outlook. Thus, it is important to continue to weigh the pros and cons of the arguments against controlled education experiments.

for international comparative understanding of what education is and how it should be judged. Researchers should capitalize on what can be learned from a natural worldwide laboratory of education systems, which over generations have come to embody many important differences in opportunities to learn, students' desire to learn, curricular structures, and teaching practices. Addressing these differences involves making inferences about what different cultures and organizations around the world consider to be excellent outcomes and excellent practice. Education policy makers should focus on success wherever they find it and could also learn from what does not seem to work. It is especially important to seek data over time in order to examine trends and the impact of changing contextual conditions. In addition, the collection and sharing of the artifacts of education (e.g., textbooks, children's writings, teachers' journals, and videotapes of classrooms) should be encouraged so that discussions of education in other places and at other times can give the users of such artifacts a better sense for the social texture and cultural intricacies that are so central to knowledge of teaching, learning, and human development.

The balance of this section focuses on the two broad types of international comparative knowledge about education; each type satisfies different needs for information. First is the knowledge needed to address issues of policy and practice currently on the public agenda. Approaches in this category focus on problems for which researchers have arrived at a working definition and for which they seek alternative solutions to compare and evaluate. Second is knowledge without reference to a particular current problem or issue—such knowledge is gathered to enhance general understanding in the long term.

Responding to Particular Policy Needs

An important purpose of cross-national research is to respond to particular policy needs. For example, an issue of particular importance to U.S. policy makers at present is whether all our students have been given adequate opportunities, appropriate standards, and justifiable expectations for learning, relative to what is done in other countries. In this respect, it is important to examine out-of-school opportunities as well as

those provided by schools. Questions range from "What can be expected of our best students?" to "What can be done about students who are poorly prepared and poorly motivated?" In short, what can be done to make sure that all students truly have adequate opportunities to learn?

Such policy issues can be addressed in three ways:

- Tracking progress within and among nations;
- Comparing the quality of education offered while explicating and attempting to understand its context; and
- Initating policy-driven comparative studies of education institutions and practices.

Tracking Progress

The lack of an adequate system of education indicators to inform education policy making has become increasingly apparent. Data are not collected regularly, systematically, or with enough coordination either to satisfy natural curiosity about education systems around the world or to answer the questions of researchers and policy makers about changes over time in education in a variety of countries. Trend data are needed on many aspects of education. We illustrate with three types of trend data that are needed:

What children in other countries are learning. International comparisons provide comparisons beyond the limits of U.S. national experience. For example, recent studies have shown that students in other countries study some mathematics topics at an earlier age than students in the United States and that consequently it may be feasible to modify U.S. curricula accordingly. In the absence of absolute standards for education, comparisons help policy makers set realistic standards. Periodic collection of data from a diverse group of countries can serve both descriptive and monitoring purposes (Bradburn and Gilford, 1990).

Factors that may be related to school achievement. Information from different countries on policy-manipulable variables that are associated with student performance can identify potential policy improvements. Such variables include the time that children in other countries spend studying important subject matters

such as mathematics, science, mother tongue, history, and foreign languages; the knowledge and expertise teachers bring to their work; and classroom size.

School finance data related to the delivery of education. Education expenditures such as teacher salaries, expenditures for in-service and other continuing education for teachers, expenditures on textbooks and other instructional materials, expenditures for special services, and class size (a determinant of the total expenditure for teacher salaries) can reflect various education policies. Education policy makers find it useful to compare trends in these indicators for the United States (a reflection of U.S. education policy) with trends in these expenditures for various countries that have education policies different from those in the United States, although the data must be used with caution because of the difficulties in obtaining comparable financial data.

Not only should national averages and other measures of central tendency be investigated, but also indices and graphic representations of variation within and between countries and trends over time as they become available.

International agencies have begun to provide information to answer questions about some aspects of education. For the past three decades a nongovernmental consortium known as the International Association for the Evaluation of Educational Achievement (IEA) has endeavored to fill a major gap by providing cross-national data on educational achievement, augmented by extensive information concerning the curriculum, classroom process, and other contextual information necessary to its interpretation. IEA studies have covered a broad range of topics: mathematics, science, reading literacy, written composition, literature, English as a foreign language, French as a foreign language, civic education, classroom environment, computers in education, and early childhood care and education. IEA has not identified or attempted to deal with all the important issues that should be addressed by indicators, nor has it been sufficiently well financed or well organized to set up regularly scheduled international assessments in all important school subject matters. Recently, however, IEA has released plans for a schedule

of 4-year cycles of studies that provide for reports: a cycle for mathematics and science with main data collection in 1995, a language cycle starting in 1996 and alternating between second or foreign language and reading literacy, a cycle starting in 1997 alternating between civics and the arts, and a cycle of special studies such as preprimary education or computers in education starting in 1998.

The Organization for Economic Cooperation and Development (OECD) has organized an extensive effort, growing out of preparatory conferences held in 1987 and 1988, to plan for the development of comparable education indicators across its member countries. Six networks were formed to deal with student achievement, education and labor market participation, features of school systems, attitudes and expectations, student flows, and costs and resources. Participation in these networks ranges between 6 and 20 countries. In 1992, OECD published a first set of international education indicators (OECD, 1992) developed by the Centre for Educational Research and Innovation (CERI). It provides comparative information for the OECD countries for three clusters of indicators: the demographic, economic, and social context of education systems; costs, resources, and school processes; and outcomes of education. Future publications of OECD indicator data, which are expected to be published at about 16-month intervals, will establish important trend information on OECD education systems for use by policy makers.

The United Nations Educational, Scientific and Cultural Organization (UNESCO) also collects country data on key aspects of education in over 160 countries. During the 1950s and 1960s, UNESCO issued a periodic publication, *A World Survey of Education*, which provided statistics on the state of education throughout the world. In 1992, UNESCO published *World Education Report 1991* (UNESCO, 1992), the first in a new biennial series that will present major trends and current policy issues in education. Focusing on basic education, the 1991 report describes world education growth since 1970 and includes information on demographic trends, adult literacy, participation in formal education, teachers, and finance. The report highlights basic education and continuing challenges in that area. Looking to-

ward the future, the report identifies two major issues: teachers and teaching (e.g., the need for additional teachers, recruitment, and training) and the assessment of students' learning achievement (e.g., school-based assessment, changes in assessment practices). The report includes education indicators for the years 1980 and 1988, thereby providing measures of change in education characteristics, for over 160 countries on a number of topics: population and gross national product; literacy, communications, and media; school entry and participation; enrollment ratios and internal efficiency in first-level education; enrollment ratios in second level education; teaching staff in preprimary, first- and second-level education; enrollment ratios and number of teachers and students in third-level education, as well as students by broad field of study for 1988 only; public expenditure on education and private enrollment; and current public expenditure on education.

These OECD and UNESCO reports are welcome additions to education information. In planning for future indicator sets, a 1991 report by the Special Study Panel on Education Indicators for the U.S. National Center for Education Statistics (NCES, 1992) is particularly relevant. The five principles advocated in this report and listed below were developed for national indicator sets, but they could be applied to cross-national sets as well:

- Indicators should address enduring issues. Although a model for indicator development oriented to education goals (i.e., define goals, specify objectives related to the goals, and develop indicators to measure progress toward the objectives and goals) could be used, it would lose its value over time as goals change. Researchers should assess what they think is important, not settle only for what they can measure.
- The public's understanding of education can be improved by high-quality, reliable indicators.
- An effective indicator system must monitor education outcomes and processes wherever they occur (not just in schools).
- An indicator system built solely around achievement tests will mislead the people in the United States.
- An indicator system must respect the complexity of the

education process and the internal operations of schools and colleges (National Center for Education Statistics, 1991:19-21).

Moreover, for use in informing policy and practice, sets of indicators are more useful than individual indicators, and these sets must be developed for a specific context, form, and level of education.

Comparing Quality While Understanding Context

The interested layperson naturally wants to know how the quality of education in the United States compares with that of other countries. Often this concern is addressed in ways that pay scant attention to the complexities of education and the particular contexts in which children are educated. Cross-national studies in education have attempted to deal with this dilemma by developing methods and measures that avoid comparing the incomparable. Thus, for example, the concepts and measures developed for IEA studies include: the intended curriculum (what the curriculum specifies), the implemented curriculum (what the teacher actually taught), and the attained curriculum (the student performance on a test of content). This refinement represents one way to achieve more comparability in the explanation of achievement outcomes (although much more improvement is called for in the implementation of these curriculum measures). In general, a great deal of additional work on explanatory variables is needed to increase the comparability of education processes and outcomes.

Still another approach to dealing with the difficulties of comparability is illustrated by earlier IEA research that focused on how to adjust results to take account of more and less selective secondary school systems to avoid comparing the elites of one system with a larger mass of students in another. The goal of these studies was to help policy makers consider the advantages and disadvantages of a more or less comprehensive secondary school system. The findings were used to support policy changes making a number of secondary school systems more comprehensive (e.g., Marklund, 1989; Bathory, 1989).

Curriculum and selectivity differences are not the only issues in understanding the context for cross-national compari-

sons. Similar issues include differences in language spoken, religion, laws, implements (e.g., tools, utensils, instruments) used, and values held, as well as attitudes toward testing that might influence motivation. Differences in cultural context can affect what is taught and when it is taught. Research is needed on the effects of these cultural differences on student achievement and the development of a theory explaining the contextual difference among nations.

Policy-Driven Comparative Studies of Educational Institutions and Practices

U.S. national education goals for the 1990s have provided strong motivation for continued interest and discussion of cross-national comparisons in mathematics and science (U.S. Department of Education, 1991). Additional motivation has been provided by the Goals 2000: Educate America Act (introduced in the U.S. House of Representatives and the U.S. Senate), which formally establishes the education goals and calls for a National Education Goals Panel and a National Education Standards and Improvement Council, which would develop model standards for student performance, curriculum, and school quality (H.R. 1804 and S.R. 846, 1993). Several implicit causal assumptions underlie public and political interest in cross-national studies of mathematics and science: (1) increased and improved curricular emphasis on mathematics or science in precollegiate education for all students will result in greater achievement for a greater number of students; (2) this will result in a greater number of better prepared students entering the natural sciences and engineering; and (3) this will in turn lead to a greater number of more productive scientists and engineers in the labor force. A final assumption is that more and better scientists and engineers will increase economic growth or productivity. Each of these premises has its critics, and each could be more systematically evaluated. At a minimum, conferences should be organized to ensure that all reasonably well-informed parties with opinions on such issues are heard, reports should be issued, and, if possible, strategies should be developed to keep the contending parties from speaking past or ignoring one another's point of view.

However, a focus on mathematics and science is not sufficient; other subjects are also important. It is critical to ensure that the complexities of education, including its variation across contexts and multiple content areas, are taken into account. Virtually every policy issue can be illuminated through study of how policy and practice vary among countries. For example, recent cross-national comparisons highlight gender differences and similarities at both the individual and societal levels of analysis. These studies can facilitate our understanding of the implications of the gender gap in academic achievement (e.g., they allow us empirically to evaluate arguments that suggest that more centralized education systems are more equitable and/or more efficient). These studies also help us understand the effects on economic growth of the expanded access of women to varying levels of education. Data from the studies can be used to evaluate arguments that suggest that the undereducation of women negatively affects economic growth (King, 1990).

The range of issues ripe for comparative inquiry includes questions of finance, decentralization versus centralization, the professionalization of teachers, the need for formal programs of teacher education, the effects of education on productivity, the effectiveness of measures whose goal is to enhance education opportunities for minority populations, the value of bilingual education, differing mixes of public and private schools (including religious schools), and even the feasibility of emphasizing moral education in public schools. Although much has been said about the world's constituting a natural laboratory, in fact the availability to U.S. audiences of current, relevant, and in-depth information on policy issues in other countries remains far too limited. Although it is true that studying education in other countries will not tell us what we *ought* to do, such study will help us to become clearer about what we value and the policy options for putting these values into practice.[3]

What does it mean to pinpoint areas in which cross-na-

[3]It should be noted that countries with decentralized education systems such as the United States can be considered to be a self-contained natural laboratory. Cross-state studies also have the potential to contribute to knowledge about the impact of education practices and policies on education outcomes.

tional differences are important to education policy and practice and to attempt to provide reasonably persuasive explanations for such differences? Heretofore, this goal has meant an attempt in cross-national surveys to estimate how much education outcomes are affected by factors controlled through education policy. This effort has been motivated by the idea that comparisons of achievement levels are not meaningful unless one can, first, identify the educational inputs and processes that contribute to observed outcome differences between countries; second, make some estimate of the contribution of each educational input to realized outcome levels, and third, consider how these effects vary by context.

But surveys by themselves are not adequate. They have focused too much on narrow measures of achievement, neglecting other important outcomes. They are not very sensitive to contextual differences, and as a result it is often very difficult to give plausible explanations for observed differences. The complexities and difficulties of acquiring a more adequate understanding of education in other countries and its possible implications for the United States can be illustrated by looking at Japan. Since progress has been made in U.S. understanding of Japanese education, let us look carefully at how this understanding has evolved, at what has been learned, and at the gaps and ambiguities that remain in understanding. Turning back to the increasing exposure to Japanese education over the past 20 years, it becomes clear that there are different phases, which have gradually led to more, though still inadequate, comparative understanding.

During the 1980s, policy makers became aware that achievement levels on conventional multiple-choice tests in mathematics and science were higher for Japanese than for U.S. students at certain grade levels. For example, the IEA Second International Mathematics Study, which gathered data in the 1981-82 school year, reported that Japanese eighth-grade students consistently scored higher than students in other countries in arithmetic, algebra, geometry, and measurement, whereas U.S. students scored in the middle of the group of countries on arithmetic and algebra and below the international average on geometry and measurement (McKnight et al., 1987:20-22). In science, in

grades 5 and 9, Japanese students also scored at the top of the group, although at grade 12 their scores were lower (International Association for the Evaluation of Educational Achievement, 1988:7, 26-27, 32-36, 48-54). This awareness was accompanied by debates over the significance of and reasons for such results. There was much disagreement over whether these measures were meaningful when taken in isolation from other evidence about education outcomes. These debates stimulated public interest, yet they were initially largely uninformed by further empirical research on these issues.

Gradually, however, there have been more focused and data-based studies that were intended to provide more in-depth understanding of Japanese practices (e.g., studies by Cummings, Peak, Rohlen, Shimahara, Stevenson, Tobin, and others). As a result, relatively uninformed endorsement or rejection of Japanese practices has partially yielded to attempts to explain these practices more thoroughly within the Japanese cultural context. More attention has been directed to the many interactive factors that might influence a particular practice or outcome, such as the amount of time spent on a task or on homework, teacher quality, and parental beliefs about the relative importance of factors influencing success in school.

As such studies become more generally known, non-comparativists can be expected to cite these findings more frequently and to incorporate them in a more general understanding of mathematics education. Perhaps in the near future there will even be a sizable market for publications about how mathematics and science are taught in Japan as well as other topics.

To the extent that such writings bear fruit in explicating practices that are effective in Japan, there will inevitably be more debate about whether similar effects might be anticipated in the United States. Although the borrowing of educational policies and practices is much criticized for infringing on cultural values and indigenous practices and for its undeniable risks of unintended consequences, it can also positively influence classroom practices. It has been and will no doubt remain widely attempted in education.

The same principles, in many respects, apply to other countries and subject areas. However, it is very difficult to carry

out multivariate analysis of natural variation that is reliable and valid. For example, the high scores of Japanese eighth-grade students relative to those of the U.S. students in the IEA Second International Mathematics Study raised questions about the quality of precollege mathematics teachers in the United States. But, although poor student outcomes may be correlated with teacher quality in terms of training characteristics or classroom method, the cause may be due to factors unrelated to teacher quality, such as the curriculum, time spent on mathematics topics in school, and the influence of the home environment. The effects of all these factors on student achievement need to be understood before we understand the relative performance of Japanese and U.S. students. Although international studies have provided detailed information on student achievement in various subjects in many countries, the causal analyses of the comparative results have been inadequate. Advances in survey analytic techniques, such as hierarchical linear modeling, offer promise of addressing those weaknesses.

Long-Term Development of Knowledge

A second important purpose of cross-national research is the long-term development of knowledge. Scholars often gather knowledge for its own sake, sometimes to expand the knowledge base in their field and at other times to develop their capacity for grasping trends and advancing ideas. In the long run, this exploration enriches and expands their understanding of the world and its complexities. They make an intellectual investment, to be drawn on later, as they interpret their research findings and the results of other related investigations.

To underwrite such investments, it is important to collect cross-national data at societal levels over reasonably long periods of time. Such data facilitate the identification of worldwide, regional, and national trends and permit analysis of the sources and effects of cross-national variation in education organization, policy, and practice. Systematic cross-national inquiry, motivated by theory and carried out over time, is needed to discourage parochial folklore and inappropriate inferences

rooted in a spatially and temporally limited vision of what education is as well as what it could or should be.

More specifically, long-term development of knowledge can affect comparative inquiry in a number of important ways:

- Becoming more cumulative and theoretical by providing synthetic works and bringing about consensus on a cross-national system of indicators, which in turn will illuminate changes across time;
- Highlighting areas in which cross-national differences are important to education policy and practice as well as provide reasonably persuasive explanations for such differences;
- Stimulating the development of a comparative dimension within all major educational specialties (e.g., empirically based comparative education has had far too little to say about the realities of classroom practice);
- Communicating the state of comparative education knowledge more effectively than at present to nonspecialists so that the thinking and discourse of educators in other fields will, as a matter of course, be informed by international comparative studies;
- Drawing attention to potential weaknesses or strengths of education systems;
- Identifying models or practices of education in other countries that have rarely, if ever, been used in U.S. education; and
- Expanding the categories and solutions that researchers consider when they think about education problems and thereby contribute to the contextual understanding of education.

In designing and carrying out such long-term efforts, a better understanding will be gained not only of contextual variation in education policies, institutions, and practices and their effects on learners, but also of more fundamental aspects of human and social development, including:

- The developmental processes of children, youth, and adults and the implications of these processes for learning;
- Education issues related to gender, race, and social class;

- The importance of cultural values in shaping the identities for which children are educated; and
- The role of education in relation to subsequent social and economic stratification in adult life.

Knowledge of a nation's education system will reveal what a nation values in its people. Is education for some or for all? What cultural norms does the education system embody? Are assumptions made that some people are unworthy of education investment? What are the personal and national payoffs of formal schooling? Who are the intended beneficiaries of the system, in addition to students? Such issues are of paramount importance in considering how nations determine who is to be educated, what is to be learned, and how all this can be accomplished. The central role of teachers in the success of an education program should be studied, as should issues concerning their recruitment, retention, training, rewarding, development, and deployment. Equally important are issues related to formal and informal sources of control and autonomy.

HOW CAN INTERNATIONAL COMPARATIVE STUDIES BE IMPROVED?

It is important to conserve and accumulate whatever knowledge and experience is needed for the betterment of future comparative international studies. In particular, since advancing the science and technology of cross-national comparative education research is a legitimate end in itself, both theoretical and empirical research aimed directly at methodological improvement should be encouraged.

To improve the way such knowledge is gathered, interpreted, and used, a number of topics that require attention are discussed in the sections that follow: research design, data analysis, and dissemination; methods for assessing current and desired outcomes of education; comparability across nations and ways of interpreting differences within varying contexts; use of ethnographic and historical studies to strengthen investigations using statistical analysis and to provide in general a deeper, richer sense of what education is, can, and should be; quality

control and monitoring; and accumulation, synthesis, dissemination, and use of cross-national knowledge about education.

Research Design, Analysis, and Dissemination

Current approaches to the design and analysis of international studies are inadequate to show how much various educational inputs and processes contribute to students' learning and later occupational endeavors and to estimate the effectiveness of these inputs and processes in different environments. In such analyses, it remains important to find better ways to take into account factors not under the control of education decision makers, such as students' family background and neighborhood setting. More generally, there is a need to be thoughtful and explicit about the underlying theories for comparative education research.

It is also important in both the design and dissemination phases of a study, to consider researchers who may wish to carry out secondary analyses of the data. Secondary analyses of large education studies frequently make major contributions to the literature and to policy concerns. For example, using data from the public-access data base developed at the University of Illinois at Urbana-Champaign for the Second International Mathematics Study (SIMS), Westbury (1992) compared mathematics achievement in the United States and Japan. He examined curriculum and achievement in grade 7/8 algebra and grade 12 elementary functions and analysis (calculus). He found that the curricula in the United States are not as well matched to the SIMS tests as are the curricula of Japan, a situation that results in lower achievement in the United States. In grade 8 algebra classes, however, the U.S. curricula are comparable to the "curriculum" of the test and the Japanese curriculum, and U.S. achievement is similar to that of Japan. His finding that the difference between Japanese and U.S. achievement is a consequence of different curricula is very different from earlier analyses of SIMS, which concluded that U.S. students performed poorly in every grade and in every aspect of mathematics tested when compared with students in the other 19 countries in the study (Crosswhite et al., 1986). He concludes that analysis focused on the undifferentiated variable of

"country" as a unit of analysis, which has been used in studies like those of the IEA or the IAEP, needs to be rethought to provide more emphasis "on analytic variables defining the properties of school systems that are common across countries but that might be distributed in different ways in different places" (p. 23).

Therefore, in designing studies, consideration should be given to fostering ways to increase the opportunities for the research community to do secondary analyses by (1) using well-established methods of analysis or (2) clarifying methodological innovations to enhance their understanding and use. It is also crucial for comparative data to be made available to the education research community in a timely and effective manner, since most researchers may not be affiliated with the organization collecting the comparative data and so will not have early access to the data. It should be possible to establish a data bank in which all published data from international education studies are entered and to make the data bank widely available to anyone with a computer and a modem.

It is equally important to make available to a wide audience discussions in nontechnical terms of the primary research questions addressed by cross-national studies. These presentations should spell out the advantages and disadvantages of alternative designs to provide answers to the research and policy questions posed. The topics covered in such a public discussion should also include the validity of the measures used in the designs and the possibilities for including outcomes other than student achievement. In clarifying issues that have arisen in previous studies, it is important to distinguish between the objectives and value of cross-sectional studies and longitudinal studies and to clarify the advantages and disadvantages of examining cumulative achievement levels over examining change in achievement levels within a given time period (such as a school year). Publishing methodologically oriented reports that have a common theme, namely the design and analysis of cross-national educational studies, or sponsoring training institutes on this subject, would contribute both to better understanding of their value to the countries concerned and to influencing the design of future studies.

Methods for Assessing Current and Desired Outcomes of Education

The research questions investigated in a country are often shaped by the methods used. Just as the global community provides a natural laboratory for comparing different education goals, systems, and methods, so it provides a variety of traditions and approaches to educational measurement. In the United States, since researchers have historically made much use of standardized multiple-choice tests of cognitive outcomes, research has emphasized ways to improve the kinds of learning that can be measured in that way. Some European countries have made greater use of essay examinations or performance assessments and have established research traditions that have placed less emphasis on reliability and objectivity as defined in the United States. In still other parts of the world, education scholars and policy makers have placed more emphasis on less tangible schooling outcomes, including personal and social values, character, and the ability to communicate or cooperate.

Cross-national studies are under pressure to draw on this plurality of methods and traditions, and at the same time they have been challenged to discover ways to quantify those kinds of learning for which good measures have not yet been found. Specific areas in need of investigation include: (1) reliable and valid performance-based measurement of cognitive learning outcomes across subject areas and grade levels; (2) measurement of personal values and other affective goals of schooling; (3) measurement of the abilities to communicate and cooperate as well as other schooling outcomes that are manifested only in a social context; (4) measurement of context variables at the level of class, school, neighborhood, and society, rather than the individual student; (5) measurement of schooling processes, including in particular a student's opportunity to learn not only content knowledge but also strategies for learning.

In planning a cross-national study, consideration should be given to whether it will serve its intended purpose. In particular, evaluation is required of its intended and likely unintended social consequences. Although it may not be possible to identify all the social consequences of an assessment, Messick (1988)

has suggested contrasting the potential social consequences of a proposed assessment with those of alternative procedures, including not testing at all. This type of appraisal contributes to a consequential basis of test validity and should preclude some of the problems in the use of international data.

Comparability Across Nations

In making cross-cultural comparisons, in addition to value and cultural issues, both logical and measurement problems arise in defining and implementing fair and meaningful practices. Two subcategories are important: first, primarily technical questions, for which different languages and practices complicate the problem of producing valid comparative information but do not, in principle, complicate the definition of sensible comparison groups or of the variable to be measured. As an example of a primarily technical problem, consider the comparison of per-pupil expenditures. Different education funding policies, accounting conventions, and currencies may complicate the development of comparable cross-national statistics. Economists have spent a large amount of time on this problem and have addressed it in great detail. Following the lead of quantitative economists, it should be possible to reach consensus on what should or should not be included as an education expenditure. In resolving such questions, useful models from other fields may include such international classification codes as the Standard International Trade Classification and the International Standard Industrial Classification of All Economic Activities. Effort should be directed to further development of such codes with a common language to describe and discuss educational organizations, processes, and outcomes so that cross-national measures and analyses of these elements can be improved.

At a second level, there are deeper substantive questions, for which different languages and practices may lead to incommensurabilities, making it difficult even to *conceptualize* variables representing corresponding attributes of students, practices, and institutions in different parts of the world. A problem in this category is the comparison of writing proficiency across languages. Problems of translations per se are

difficult enough, but, beyond translation, languages have different conventions for organizing prose and presenting ideas.

Comparisons among countries can be difficult to analyze because of the large number of differences in the countries being compared. The large number of interactions between variables makes it difficult to identify which variables influence outcomes and which are covarying with those that influence the outcomes. Both the problem of covariation and the problem of interaction cause difficulties in making interpretations in comparative studies. These difficulties in drawing conclusions from comparisons do not mean that comparisons should be ignored—there are many useful advantages afforded by such investigations as discussed earlier in this report.

In cross-national research, such issues of what constitutes an intelligent comparison could be more readily addressed if some studies were focused on specific topics within a few participating countries instead of omnibus studies welcoming as many countries as wish to join. For example, limiting the countries in a study to a few developed countries in Europe and North America (or to groups of developing countries, African countries, Asian countries, or South American countries) would reduce the number of differences between countries being compared, making it easier to identify which variables influence outcomes. Conceptualizing variables representing attributes of students, practices, and institutions would also be easier for a limited study than for a worldwide study. Finally, if the study were focused on a few topics, for example, teaching practices and student achievement in secondary school chemistry, the topics could be explored in greater depth than in a study with countries having very disparate education systems.

In addition, a limited study would have fewer problems with languages, culture, and value systems; would have fewer communication problems in study administration; would require less training; would be less costly; and could be completed on a more timely basis. Such a study could also be designed to meet the policy needs of the participating countries.

Qualitative Studies and Large-Scale Surveys

Many significant questions in comparative education are best addressed by small, focused studies, which may draw on a broad range of techniques and provide a deeper, richer sense of what education is, can, and should be. Thus, in addition to large-scale surveys, there is a need for a wide range of other cross-national research, such as ethnographic studies, case studies, small-scale focused quantitative and qualitative studies, and historical studies that would allow us to understand what it means to be educated in diverse settings around the world. Such studies go beyond the exploratory and the descriptive. They have become essential parts of the explanatory repertoire. These studies provide ways of analyzing and explaining a variety of processes, conditions, and contexts. They help to uncover patterns of interaction and to interpret complex situations both in the classroom and in the larger community. There is a great need for small, in-depth studies of local situations that would permit cross-cultural comparisons capable of identifying the myriad of causal variables that are not recognized in large-scale surveys. In fact, much survey data would remain difficult to interpret and explain without the deep understanding of society that other kinds of studies provide. Given that research in cross-national contexts benefits from increased documentation of related contextual information, it would be useful to combine large-scale surveys and qualitative methods.

Ethnographic and other qualitative studies are especially important in clarifying the perspectives of many diverse groups of learners and their teachers. Such groups include gifted and talented students, individuals with disabilities, racial and ethnic minorities, women, and religious groups that reject the state's secular systems of education.

But there is also a role for the large-scale surveys. About the only way to obtain a simple numerical comparison of a large number of countries on a common set of measures is with a large-scale survey. Large-scale studies also permit curriculum analysis on a scale that could not possibly be done adequately by a few independent researchers, because such studies require a high degree of international organization and structure. For example, the preliminary curriculum analyses

of data collected by IEA in the Third International Mathematics and Science Study (TIMSS) (Schmidt, 1993) reveal a large and interesting variation in mathematics and science textbooks across a large number of countries—variation that can be helpful in better understanding the role of the teacher in various countries. Textbooks can have a major influence on teachers' curricular decisions (Schmidt et al., 1987). Teachers are more likely to teach the topics included in the textbook than others not included. Large-scale studies such as TIMSS or the OECD indicators project will also provide trend data, and independent researchers are unlikely to have the commitment, longevity, or resources to produce such data in adequate fashion.

Improving Quality Control and Monitoring

Due to the lack of continuity in funding and personnel as well as the lack of an organizational structure capable of maintaining rigorous quality control, the overall quality of some large-scale cross-national studies of educational achievement has suffered. In many countries, the quality of the sampling, the translation of instruments into the national language(s), and the collection and management of data have been impeccable, but in other countries errors have been made in one or more of these processes. The results have been damaging in two respects: on one hand, suspect conclusions have been accepted, and on the other hand, critics of large-scale studies have overgeneralized and exaggerated these faults and asserted that nothing in the whole enterprise is valid or to be believed. For example, whereas in some countries, exclusions from target populations may have been substantial enough to mislead the user about the effects of a given country's schooling policies and practices, such exclusions can be clearly identified in reports and, in many cases, shown not to be so extensive as to affect conclusions drawn about the country as a whole.

Hence increased attention to controlling all sources of errors is important. This includes such matters as defining comparative populations, constructing sampling frames and selecting samples, developing instruments and maintaining cross-national equivalency of instrumentation, administering instruments and recording data, entering and editing data, coding

and scoring, and weighting and analysis. Documentation on the steps taken to control errors is essential at all stages: at the planning stage, during data collection as a guide to study personnel in all locations, and as a necessary part of reporting what was done.

Fostering Accumulation, Synthesis, and Dissemination

Dissemination and utilization of international education data remain weak links in the information-reform connection within education. Public discussion of cross-national education research is frequently superficial or incorrect or both. The use of international data in the United States needs close scrutiny. In the past, test results were often extracted from a study and reported independently of other variables that provide essential context for the results. Sometimes the limitations, such as sampling problems, were glossed over as if they did not matter, and at other times they were exaggerated as if they were sufficient to rebut everything that could be learned from such studies. In the future, researchers and organizations responsible for cross-national studies need to begin early in project cycles to plan for public and professional discussion of the results and to improve the presentation of findings to all concerned. Press conferences should be planned that respond to the needs of journalists and at the same time present a more in-depth background on what has been done. These meetings should clarify the importance of considering variables such as demographics, expenditures, and teacher training in conjunction with achievement scores in presenting the results of the increasingly sophisticated research on education that is under way or already available.

In addition to a better understanding of how and what to communicate, consideration is needed of the potential contributions of new communication technologies and data bases to the planning of high-quality cross-national studies and to the dissemination and utilization of results. The rapidity and simplicity of communication by facsimile and by electronic networks (e.g., BITNET and Internet) have revolutionized international research in the last five years. As cross-national studies produce items and instruments, it will be important to con-

sider how all the information generated by such studies might be made more accessible through electronic communication networks.

The presentation and interpretation of current results could be improved if there were an easily accessible repository of cross-national research. Students of education—from undergraduate prospective teachers to senior researchers in highly specialized subfields—ought to have readily at hand a literature (good research published in reasonable places) and associated data bases that would give them access not just to a description of the structure of other education systems, but to an understanding of how such systems work, with particular emphasis on how context influences school organization, teaching practices, and learning outcomes.

A CHALLENGE TO REALIZE THE VISION AND A POSSIBLE APPROACH

The introduction described a vision of what international comparative research and data collection might produce in the next 20 years. It is questionable, however, whether existing organizations within the United States or international organizations possess the full infrastructure required to accomplish the vision that we have projected.

Improving the Infrastructure for Cross-National Research

There is a need either to expand the role of existing organizations for cross-national research in education or to create new ones. The ideal would be something similar to what was envisioned originally in the late 1940s for the education sector of UNESCO—a forum in which the best ideas on education could be researched and shared, in which jointly financed projects could be identified and appraised, and in which professionals from many disciplines could discuss results, publish freely, and assess the strengths and weaknesses of such work. UNESCO did not carry out this mandate because many member states placed a low priority on the importance of education research, a situation that severely limited the fiscal and personnel resources available for the activity.

In short, what is needed now for education is similar in many respects to what is already available in epidemiologic research through the World Health Organization (WHO); demographic research through the United Nations Population Fund (UNFPA); and food plant research through the Consultative Group for International Agricultural Research (CGIAR). Prerequisite to achieving such an infrastructure would be to diversify the institutional relationships beyond those that now typify education research and, in particular, to incorporate the views of not only recognized scholars in diverse disciplines but also such parties as business and labor leaders, educational managers and teachers, and a wide range of political authorities.

Nongovernmental agencies comprise a critical segment of this infrastructure. Priority should be given to developing centers of excellence, whether in universities or elsewhere, that are independent of policy imperatives. Such centers, given appropriate support, could be a catalyst for identifying and training people with the right skills and sensitivities. Ultimately it would be desirable to have a network of education research centers that would enable scholars from around the world to undertake, more readily than is now the case, cross-national studies in societies other than their own.

A Possible Approach

Since the task of establishing a more coherent and effective worldwide system for cross-national research and data collection in education will be neither straightforward nor easy, it must be shared among many organizations in many countries. One possible approach would begin with consultations and alliances among stakeholders and other interested parties; first steps in this direction are the two conferences described in the preface. These consultation meetings should result in a report, perhaps issued jointly by several organizations or agencies, on what a better infrastructure to conduct cross-national studies in education might look like, what it might cost, and what it might produce. Before its release, ideas from the report should be shared with senior policy makers whose interests in education are known. Discussion should take place at ministerial

level, since government ministers clearly do have interest in these matters, especially the matter of policies, and if possible should include heads of government. The purpose of the discussion would be to interest the ministers in building the capacity for meeting these information needs.

The ultimate goal to be pursued through consultation and the formation of alliances would be the establishment, maintenance, and improvement of networks, organizations, and other institutional mechanisms internationally, within the United States, and in other nations for funding, managing, conducting, and disseminating cross-national educational studies. Sound cross-national research and data collection require stable institutional mechanisms, both to secure the cooperation of different nations and to ensure adequate, reliable funding for technically sound planning, execution, analysis, interpretation, and dissemination.

These mechanisms must also be designed in such a way that the education community identifies and prioritizes the research issues, questions, and topics to be investigated. The infrastructure would provide a forum for education researchers to deliberate research agendas and priorities, as well as the objectives of specific studies. Priority projects could be selected from any of the three categories of projects described earlier in this report:

- Projects to develop and produce education indicators;
- Comparative education assessments (with associated explanatory research and curriculum analyses) that are better organized, with more justifiable and precise objectives, more valid designs and measures, more rigorous quality control and more adequate funding, and conducted on a regular basis; and
- Other types of studies that can help interpret much of the variation revealed through education indicators and cross-national assessment data, as well as investigations of other comparative questions and issues.

The agendas would serve both of the purposes for international educational research identified earlier in this paper: (1) long-term development of knowledge and (2) knowledge for

policy makers. Knowledge for policy makers can be derived either from studies to respond to particular policy needs for data or from policy-relevant data not tied to any specific current issue (e.g., data provided by education indicator systems featuring regular, recurrent data collections to permit analysis of trends). Education researchers can, do, and should respond to information needs connected to current policy questions. A sound infrastructure would also provide a mechanism for channeling and responding to policy-relevant questions.

In summary, the process of establishing a more coherent and effective worldwide system for cross-national research and data collection in education will have to be one that takes into account the views of interested parties, the technical and substantive state of the art, and the promise of cross-national studies for responding to policy needs and advancing knowledge.

REFERENCES

Bathory, Zoltan

1989 The Hungarian Experience. In Alan Purves (ed.), *International Comparisons and Educational Reform*. Alexandria, Virginia: Association for Supervision and Curriculum Development.

Bradburn, Norman M., and Dorothy M. Gilford, Editors

1990 *A Framework and Principles for International Comparative Studies in Education*. Board on International Comparative Studies in Education. Washington, D.C.: National Academy Press.

Crosswhite, F.J., J.A. Dossey, J.O. Swafford, C.C. McKnight, T.J. Conney, F.L. Downs, D.A. Grouws, and A.L. Weinzweig

1986 *Second International Mathematics Study: Detailed Report for the United States*. Champaign, Illinois: Stipes.

International Association for the Evaluation of Educational Achievement

1988 *Science Achievement in Seventeen Countries: A Preliminary Report*. New York: Pergamon Press.

King, E.M.

1990 *Educating Girls and Women: Investment in Development*. Washington, D.C.: The World Bank.

Marklund, Inger

1989 The Swedish Experience. In Alan Purves (ed.), *International Comparisons and Educational Reform*. Alexandria, Virginia: Association for Supervision and Curriculum Development.

McKnight, Curtis C., F. Joe Crosswhite, John A. Dossey, Edward Kifer, Jane O. Swafford, Kenneth J. Travers, and Thomas J. Conney

1987 *The Underachieving Curriculum: Assessing U.S. School Mathematics From An International Perspective*. Champaign, Illinois: Stipes.

Messick, Samuel

1988 The once and future issues of validity: Assessing the meaning and consequences of measurement. In H. Wainer and H.I. Braun (eds.), *Test Validity 1988*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Meyer, Michael M., and Stephen E. Fienberg, Editors

1992 *Assessing Evaluation Studies: The Case of Bilingual Education Strategies*. Panel to Review Evaluation Studies of Bilingual Education, Committee on National Statistics. Washington, D.C.: National Academy Press.

National Center for Education Statistics (NCES)

1991 Education Counts, An Indicator System to Monitor the Nation's Educational Health. Report of the Special Study Panel on Education Indicators to the Acting Commissioner of Education Statistics, U.S. Department of Education, September. National Center for Education Statistics, Washington, D.C.

Organisation for Economic Cooperation and Development (OECD)
1992 *Education at a Glance*. Organisation for Economic Cooperation and Development, Paris, France.

Schmidt, W.H.
1993 Survey of Mathematics and Science Opportunity: Preliminary Results of the Curriculum Analyses. Presentation to the Board on International Comparative Studies in Education, June 18.

Schmidt, W.H., A.C. Porter, R.E. Floden, D.J. Freeman, and J.R. Schwille
1987 Four patterns of teacher content decision making. *Journal of Curriculum Studies*. 19(5):439-455.

United Nations Educational, Scientific and Cultural Organization (UNESCO)
1991 *World Education Report 1991*. United Nations Educational, Scientific and Cultural Organization, Paris, France.

U.S. Congress, House
1993 Resolution 1804. (Goals 2000: Educate America Act) 103rd Congress, 1st Session.

U.S. Congress, Senate
1993 Resolution 846. (Goals 2000: Educate America Act) 103rd Congress, 1st Session.

U.S. Department of Education
1991 *America 2000: An Education Strategy Sourcebook*. U.S. Department of Education, Washington, D.C.

Westbury, Ian
1992 Comparing American and Japanese achievement: Is the United States really a low achiever? *Educational Researcher* 21(5)(June-July):18-24.

BIOGRAPHICAL SKETCHES

JOSEPH F. ALIBRANDI is chair and chief executive of Whittaker Corporation and chair of BioWhittaker, Inc. He is a director of BankAmerica Corporation and Bank of America NTS&A, Catellus Development Corporation, Jacobs Engineering Group, and Santa Fe Pacific Corporation. Prior to joining Whittaker, he managed the missiles systems division of Raytheon Company for 18 years, most recently as senior vice president. He is the former chair of the California Business Roundtable Task Force on Education. He also served as chair of the Twelfth District Federal Reserve Bank and chair of the International Policy Committee of the U.S. Chamber of Commerce. He is a member of the advisory board of Policy Analysis for California Education, chair of ExCEL-Choice in Education League, and director of the Institute for Contemporary Studies. He received a B.S. in mechanical engineering from the Massachusetts Institute of Technology and a D.Sc. in business administration from Bryant College.

GORDON M. AMBACH is executive director of the Council of Chief State School Officers, which represents state-level interests in education and provides advocacy for their state and national education policy positions. Previously he served for 10 years as the New York State commissioner of education and president of the University of the State of New York. He also held appointments with the U.S. Office of Education in legislation and program planning. His advisory roles include membership on commissions and panels on learning technology, job training, the arts, and education statistics and assessment. He is currently the U.S. representative to the General Assembly of the International Association for the Evaluation of Educational Achievement and a member of the Standing Committee. His professional interests have centered on education policy making and legislation. He has a B.A. from Yale University and an M.A. in teaching from Harvard University.

LLOYD BOND is professor of educational research methodology and senior scientist in the Center for Educational Research

and Evaluation at the University of North Carolina. Previously he was associate professor of psychology and senior scientist at the Learning Research and Development Center at the University of Pittsburgh. His research interests include educational and psychological measurement, test bias, and cognitive processes underlying test performance. Other recent National Research Council memberships include the Study Group on Guidelines for Mathematics Assessment, the Working Group on Science Assessment Standards, and the Committee on the Performance of Military Personnel. He received a B.S. in psychology from Hillsdale College and M.A. and Ph.D. degrees in psychology from Johns Hopkins University.

HERMAN CHERNOFF is professor of statistics at Harvard University. He has also taught statistics at the University of Illinois, Stanford University, and the Massachusetts Institute of Technology. From 1959 to 1960 he was a fellow at the Center for Advanced Study of Behavioral Sciences. His research interests are in statistical problems in econometrics, the sequential design of experiments, rational selection of decision functions, large sample theory, and pattern recognition. Chernoff's other National Research Council responsibilities have included membership on the Commission on Physical Sciences, Mathematics, and Resources and the Board on Mathematical Sciences. He has a B.S. from City College of New York and Sc.M. and Ph.D. degrees in applied mathematics from Brown University, as well as honorary Sc.D. degrees from Ohio State University and Israel Institute of Technology.

DOROTHY M. GILFORD is director of the National Research Council's Board on International Comparative Studies in Education. Formerly she served as director of the National Center for Education Statistics and as director of the mathematical sciences division of the Office of Naval Research. Her interests are in research program administration, the organization of statistical systems, and education statistics. A fellow of the American Statistical Association, she has served as vice president of the association, chair of its committee on fellows, and chair of its committee on international relations in statistics. She is a member

of the International Statistics Institute. She received B.S. and M.S. degrees in mathematics from the University of Washington.

TIMOTHY H. GOLDSMITH is Andrew W. Mellon professor of biology at Yale University. Previously he has been a junior fellow in the Society of Fellows at Harvard University and an instructor at the Marine Biological Laboratory in Woods Hole, Massachusetts. From 1967 to 1968 he was a Guggenheim fellow; he is a fellow of the American Academy of Arts and Sciences. His research interests are in neurophysiology and biochemistry of light sensitive systems, particularly the vision of arthropods and the color vision of birds. He chaired the National Research Council's Committee on High School Biology Education. He has a B.A. from Cornell University and a Ph.D. from Harvard University.

JAMES W. GUTHRIE is codirector of Policy Analysis for California Education (PACE) and professor of education, both at the University of California at Berkeley. Before joining the University of California faculty, Guthrie was education specialist for the United States Senate. His research interests include education policy, school finance and governance, and the reform of education systems. His publications include books on school finance, educational administration, strategic planning, and teacher education. He has served as vice president of the American Educational Research Association and currently edits that organization's journal, *Educational Evaluation and Policy Analysis*. He has been honored as an Alfred North Whitehead postdoctoral fellow at Harvard University, visiting fellow at the Department of Educational Studies of Oxford University, and in 1990 was named the American Education Research Association's first senior fellow. He received B.A., M.A., and Ph.D. degrees from Stanford University.

EDWARD H. HAERTEL is professor of education at Stanford University. His research interests center on psychometrics and assessment policy, especially large-scale assessments. He has worked with the National Center for Education Statistics; the

National Assessment Governing Board; the Center for Research on Evaluation, Standards, and Student Testing; and other public and private organizations on issues concerning the National Assessment of Educational Progress and with various states on issues concerning assessment programs. He has a B.A. in mathematics from the University of Wisconsin and a Ph.D. in education from the University of Chicago.

STEPHEN P. HEYNEMAN is the manager of human resources in the Technical Department for the Europe and Central Asia and Middle East and North Africa Regions of the World Bank. His responsibilities cover policies and lending strategies in health, education, and labor markets. His 1970 academic achievement survey of elementary schools in Uganda was the first in Sub-Saharan Africa, and he has been utilizing international surveys to study problems of educational quality in developing and industrialized countries since that time. His other research interests include higher education finance and management, policy reforms in vocational education, examinations and standardized testing, cognitive skills, and labor productivity. He is currently managing a study of Russian universities. He is the president of the Comparative and International Education Society. He received a B.A. in political science from the University of California at Berkeley, an M.A. in African studies from the University of California at Los Angeles, and a Ph.D. in comparative education from the University of Chicago.

DANIEL G. HORVITZ is distinguished institute scientist at the Research Triangle Institute where he has also been executive vice president and vice president for statistical sciences. His research interests are survey design and methods, with emphasis on cost-effective ways to improve data quality. He is a fellow of the American Association for the Advancement of Science and the American Statistical Association, which he has served as vice president, and a member of the International Statistical Institute. He received a B.S. from the University of Massachusetts and a Ph.D. in statistics from Iowa State University.

MICHAEL W. KIRST is professor of education and codirector of Policy Analysis for California Education (PACE), both at Stanford University. Before joining the Stanford faculty, he held several positions with the federal government, including staff director of the U.S. Senate Subcommittee on Manpower, Employment and Poverty and director of program planning and evaluation for the Bureau of Elementary and Secondary Education in the U.S. Office of Education (now the U.S. Department of Education). He is a member of the research staff of the Center for Policy Research in Education and has served as a member of the National Academy of Education (since 1979), vice president of the American Educational Research Association, commissioner of the Education Commission of the States, and member (1975-81) and president (1977-81) of the California State Board of Education. He has authored and coauthored several books on schools and public accountability. He received an A.B. from Dartmouth College and M.P.A. and Ph.D. degrees from Harvard University.

GAEA LEINHARDT is a senior scientist at the Learning Research and Development Center and professor of education at the University of Pittsburgh. She has directed projects on program evaluation, teacher assessment, and teacher expertise in mathematics and history. Her research interests focus on classroom instruction and learning in specific subject areas as well as in understanding how instructional explanations are constructed and when they are effective. She has worked on the design of innovative assessment techniques for teachers and students for the National Board of Professional Teacher Standards and several state departments of education. She received a B.A. in Russian civilization and an M.S. in urban education from the University of Chicago and a Ph.D. in research methodology from the University of Pittsburgh.

PAUL G. LEMAHIEU is director of the Delaware Education Research and Development Center and associate professor of educational studies at the University of Delaware. He also holds a senior appointment in the cabinet of the state superintendent of public instruction as special advisor for educational

policy in the Delaware Department of Public Instruction. He has served as director of research, evaluation, and student assessment for the Pittsburgh public schools and adjunct associate professor of education at the University of Pittsburgh. His research interests are in assessment, minority education, and effective schools. His other National Research Council memberships include the Mathematical Sciences Education Board and the National Summit on Mathematics Assessment Steering Committee. He received an A.B. in philosophy and psychology from Yale University, an Ed.M. in educational development and evaluation from Harvard University, and a Ph.D. in educational research methodology from the University of Pittsburgh.

MARY M. LINDQUIST is Fuller E. Callaway professor of mathematics education at Columbus College and is president of the National Council of Teachers of Mathematics and vice chair of the National Research Council's Mathematical Sciences Education Board. Her teaching has included junior high school, high school, and college mathematics. From 1972 to 1976 she was assistant scientist at the University of Wisconsin Research and Development Center. Professional activities include serving on the executive committee of the Board of Directors of the National Council of Teachers of Mathematics, on a working group for the development of the *Curriculum and Evaluation Standards for School Mathematics*, on the advisory board to the National Center for Research in Mathematical Sciences Education, and on the National Advisory Governing Board's committee to set standards for the National Assessment of Educational Progress as well as on the committees to develop items for the past six mathematical assessments. She has written publications on elementary and secondary mathematics education and served on the editorial board of the *Journal for Research in Mathematics Education*. She received a B.A. in mathematics education from the University of North Carolina, an M.A. in mathematics from Southern Methodist University, and a Ph.D. in mathematics education from the University of Wisconsin.

HAROLD NISSELSON is an independent statistical consultant in Falls Church, Virginia. He has recently served as a consultant in studies for several U.S. government agencies, the government of Norway, and the National Board for Professional Teaching Standards. Previously he has served as senior survey statistician at Westat, Inc., associate director for statistical standards and methodology at the U.S. Bureau of the Census, and assistant director for research at the National Center for Education Statistics. His other National Research Council memberships have included the Panel on Incomplete Data, the Committee on Indicators of Precollege Science and Mathematics Education, and the Panel for Science-Based Careers: A Pilot Study of the Dynamics of Undergraduate Student Choices. He received a B.S. from the City College of New York and has done further study at the Graduate School of New York University and the Graduate School of the U.S. Department of Agriculture.

ANDREW C. PORTER is director of the Wisconsin Center for Education Research and professor of educational psychology at the University of Wisconsin, Madison. Previously he was codirector of the Institute for Research on Teaching and professor of education at Michigan State University. He has also served as associate director of the National Institute of Education, U.S. Department of Education. His research interests include research on teaching education policy and psychometric research. His other National Research Council memberships include the Study Group on Guidelines for Mathematics Assessment, the Panel to Evaluate the National Center for Education Statistics, and the Committee on the Federal Role in Education Research. He received a B.S. in education from Indiana State University and M.S. and Ph.D. degrees in educational psychology from the University of Wisconsin.

FRANCISCO O. RAMIREZ is professor of education and sociology at Stanford University. He is on the editorial boards of *Comparative Education Review* and *Contemporary Sociology* and has served as associate editor of *Sociology of Education*. He is the editor or coauthor of a number of publications. His re-

search interests include studies of the origins and expansion of mass schooling; women's access to education; and education, science, and development. Ramirez received a B.A. from De La Salle College in the Philippines and M.A. and Ph.D. degrees from Stanford University.

JOHN R. SCHWILLE is professor and assistant dean for international studies at the College of Education at Michigan State University. His experience includes qualitative as well as quantitative education research in both international and U.S. domestic settings. His involvement in international assessment dates from 1972, when he received a Spencer Fellowship at Stockholm University for analysis of data collected by the International Association for the Evaluation of Educational Achievement. His current work focuses on comparative issues in school organization and teacher effectiveness. He received a B.A. in the history of France from Harvard University and a Ph.D. in comparative sociology of education from the University of Chicago.

MARSHALL S. SMITH is undersecretary of the U.S. Department of Education. He is on leave from responsibilities as dean of the School of Education and professor of education at Stanford University. During 1992-1993, he was a fellow at the Center for Advanced Study in the Behavioral Sciences. He has been an associate professor at Harvard University and a professor at the University of Wisconsin. From 1983 to 1986 he was assistant commissioner for policy studies in the U.S. Office of Education and associate director of the National Institute of Education. He was a member of the Committee on Precollege Indicators of Science and Mathematics Education of the National Research Council's Commission on Behavioral and Social Sciences and Education, chair of the Pew Forum on K-12 Education, chair of a National Goals Panel subgroup on the National Assessment System, and a member of the National Council on Education Standards and Testing. His current research interests are systemic school reform, educational equity, and compensatory education. He received an A.B. in psychology from Harvard College, an Ed.M. in measurement and sta-

tistics, and an Ed.D. in measurement and statistics from the Harvard Graduate School of Education.

FLORALINE I. STEVENS is director of the Program Evaluation and Assessment Branch for the Los Angeles Unified School District. Currently on leave from the district, she is a program director in the Education and Human Resources Directorate of the National Science Foundation. She served as the second senior research fellow at the National Center for Education Statistics. Her interests are program evaluation and research issues focused on at-risk students. She served as vice president for the division on school evaluation and program development of the American Educational Research Association (AERA), chair of the Research Steering Committee for the Council of Great City Schools, and chair of the School Research Directors at AERA; she has served on the National Indicators Panel for the U.S. Department of Education. She received a B.S. from the University of Southern California and M.Ed. and Ed.D. from the University of California at Los Angeles.

JUDITH V. TORNEY-PURTA is professor of human development and affiliate professor of psychology at the University of Maryland, specializing in applied developmental and educational psychology. Formerly she was professor of psychology at the University of Illinois at Chicago. She was the senior author of the International Association for the Evaluation of Educational Achievement volume reporting the 10-country study of civic education. She has served on the Fulbright Senior Scholar Selection Panel in Psychology and as chair of the International Relations Committee of the American Educational Research Association, and she is a fellow of the American Psychological Association. Her current research interests are in applying cognitive modeling techniques to measure the complexity and constraints in adolescents' views of the political and economic world. She received an A.B. in psychology from Stanford University and M.A. and Ph.D. degrees in human development from the University of Chicago.

NATIONAL ACADEMY PRESS
The National Academy Press was created by the National Academy of Sciences to publish the reports issued by the Academy and by the National Academy of Engineering, the Institute of Medicine, and the National Research Council, all operating under the charter granted to the National Academy of Sciences by the Congress of the United States.